

AD _____

AWARD NUMBER: W81XWH-12-2-0052

TITLE: Integrative Cardiac Health Project, Windber Research Institute

PRINCIPAL INVESTIGATOR: Darrell L. Ellsworth, PhD

RECIPIENT: Windber Research Institute
WINDBER, PA 15963

REPORT DATE: July 2014

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution is unlimited.

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE July 2014		2. REPORT TYPE Annual		3. DATES COVERED 22 Jun 2013 - 21 Jun 2014	
4. TITLE AND SUBTITLE Integrative Cardiac Health Project, Windber Research Institute				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-12-2-0052	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Darrell L. Ellsworth, PhD Email: d.ellsworth@wriwindber.org				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Windber Research Institute Windber, PA 15963				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The Integrative Cardiac Health Program (ICHP) aims to understand the complex interactions of numerous molecular components that characterize individuals or populations at risk for cardiovascular disease (CVD). We are working to identify molecular networks that define cardiovascular risk and correspond to lifestyle changes that may influence the trajectory of disease progression. Given the increasing trends of major CVD risk factors in the US military population and potentially devastating consequences on combat readiness, our efforts are directed to (1) better understanding the relationships of war stressors and CVD risk at the molecular level before onset of clinical disease, and (2) outcomes-based patient empowering lifestyle solutions to prevent disease. WRNMMC ICHP has developed a standardized and personalized lifestyle intervention program that has resulted in significant improvements in cardiovascular risk markers such as C-reactive protein (CRP), glucose, insulin, glycosylated hemoglobin, and lipids. Our previous research was the first to identify significant gene expression changes associated with an ultra-intensive lifestyle change program and to show that genetic variants at genes involved in lipid metabolism influence lipid response. We hypothesize that (1) identifying genetic influences on CVD and integrating information on dietary, behavioral, and lifestyle factors will provide important information on CVD risk reduction and (2) discovering new genes in previously associated pathways will reveal new molecular influences on cardiovascular risk reduction.					
15. SUBJECT TERMS- none listed					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	9	19b. TELEPHONE NUMBER (include area code)

Table of Contents

	<u>Page</u>
1. Introduction	4
2. Keywords	4
3. Overall Project Summary	4
4. Key Research Accomplishments	8
5. Conclusion	8
6. Publications, Abstracts, and Presentations	8
7. Inventions, Patents and Licenses	8
8. Reportable Outcomes	8
9. Other Achievements	8
10. References	8
11. Appendices	8

[SF298]

The Integrative Cardiac Health Program (ICHP) aims to understand the complex interactions of numerous molecular components that characterize individuals or populations at risk for cardiovascular disease (CVD). We are working to identify molecular networks that define cardiovascular risk and correspond to lifestyle changes that may influence the trajectory of disease progression. Given the increasing trends of major CVD risk factors in the US military population and potentially devastating consequences on combat readiness, our efforts are directed to (1) better understanding the relationships of war stressors and CVD risk at the molecular level before onset of clinical disease, and (2) outcomes-based patient empowering lifestyle solutions to prevent disease.

WRNMMC ICHP has developed a standardized and personalized lifestyle intervention program that has resulted in significant improvements in cardiovascular risk markers such as C-reactive protein (CRP), glucose, insulin, glycosylated hemoglobin, and lipids. Our previous research was the first to identify significant gene expression changes associated with an ultra-intensive lifestyle change program and to show that genetic variants at genes involved in lipid metabolism influence lipid response. We hypothesize that (1) identifying genetic influences on CVD and integrating information on dietary, behavioral, and lifestyle factors will provide important information on CVD risk reduction and (2) discovering new genes in previously associated pathways will reveal new molecular influences on cardiovascular risk reduction.

This research is using state-of-the-art next-generation DNA and RNA sequencing to address the following research questions:

1. Can inherited variants in specific genes that influence cardiovascular disease be identified through whole-genome sequencing of an appropriate population of affected individuals?
2. Does surgically-induced weight loss alter patterns of gene expression in adipose tissue and peripheral blood, and do these molecular changes have prognostic value in predicting weight loss success and improving the effectiveness of treatment programs for obesity?
3. Can changes in whole-transcriptome expression: (a) improve the effectiveness of lifestyle programs by identifying patients who are unlikely to benefit from the conventional program and should follow a customized program tailored to their individual needs; (b) be used to reduce health outcome disparities between men and women and between specific subgroups of patients; and (c) identify previously unknown molecular pathways that influence heart disease, and provide insights into cardiovascular disease development that may have important consequences for cardiac treatment programs?

1. INTRODUCTION: There is an increasing prevalence of obesity and cardiovascular disease (CVD) risk factors in the military population, which is negatively affecting operational readiness. The ability to prevent heart disease and reduce its overall impact on morbidity, would increase the quality of life among military personnel and their dependents, and has the potential to generate enormous cost savings for the DoD. In the Integrative Cardiac Health Program (IHP), we are investigating physiological and molecular responses to risk factor modification interventions in individuals or populations at risk for cardiovascular disease (CVD). We aim to better understand CVD risk at the molecular level before onset of clinical disease, and develop outcomes-based patient empowering lifestyle solutions to prevent disease. Through this research, our objectives are to (1) identify genetic influences on CVD and integrate information on dietary, behavioral, and lifestyle factors to provide important information on CVD risk reduction and (2) discover new genes in previously associated pathways to reveal new molecular influences on cardiovascular risk reduction.

2. KEYWORDS: Lifestyle modification, cardiovascular disease, obesity, gene expression, RNA sequencing, gender differences, molecular response, diet, exercise.

3. OVERALL PROJECT SUMMARY:

For all tasks, much effort was devoted to training and trouble-shooting to be able to consistently generate high quality DNA and RNA sequence data. Experiments are ongoing to enhance DNA removal from RNA. The initial training by Dr. Clifton Dalgard and colleagues from the Uniformed Services University enabled us to optimize our Illumina work-flows for library preparation and clustering for eventual running on the HiSeq. We continued experiments to optimize DNA removal from RNA and globin-reduction in total RNA samples from blood samples. (7/1/2013—6/30/2014)

We made important progress on preparing libraries, making the appropriate dilutions, and running samples on the MiSeq machine, and gained valuable experience interpreting QC results from experimental sequencing runs. (1/1/2014—6/30/2014)

Illumina service representatives recertified the HiSeq, MiSeq, and cBot machines on June 11, 2014 and the machines now have service contract coverage. All machines are currently functioning normally. An initial RNA sequencing run using 40 samples generated 138 billion bases of RNA sequence in ~1.15 billion raw reads with very good quality (99.9% accuracy) at 49x coverage. (4/1/2014—6/30/2014)

Throughout the year, blood samples were collected from 83 patient follow-up visits, and 1238 plasma aliquots were banked. (7/1/2013—3/31/2014)

R Training – Our statistician completed a 3-day course that provided individualized instruction in the basics of R computer language, which will be necessary for analyzing the large-scale sequencing data. (1/1/2014—3/31/2014)

All research protocols for this project were submitted to and approved by the Chesapeake IRB. (1/1/2014—3/31/2014)

Task #1: Epigenetic changes in DNA (genome-wide patterns of methylation) during CV risk reduction

In this task, we are examining patterns of DNA methylation across the entire genome in circulating leukocytes in response to cardiovascular risk reduction (lifestyle and surgically-assisted) using new and current participants in our ICHP programs. We will seek to identify changes in methylation in specific areas of the genome and relate these changes to known and novel genes influencing heart disease. Results from this research may be useful in further understanding molecular mechanisms associated with changes in CV risk factors and regulatory processes involved in heart disease development.

DNA was isolated from 30 whole blood samples using the Quick gDNA Blood Mini kit (Zymo Research) in the following groups: intensive lifestyle baseline (5) and one year (5), laparoscopically placed adjustable gastric banding (LAGB) baseline (5) and one year (5), control baseline (5) and one year (5). OD260/280 ratios, OD260/230 ratios, and DNA concentrations are shown in the table below.

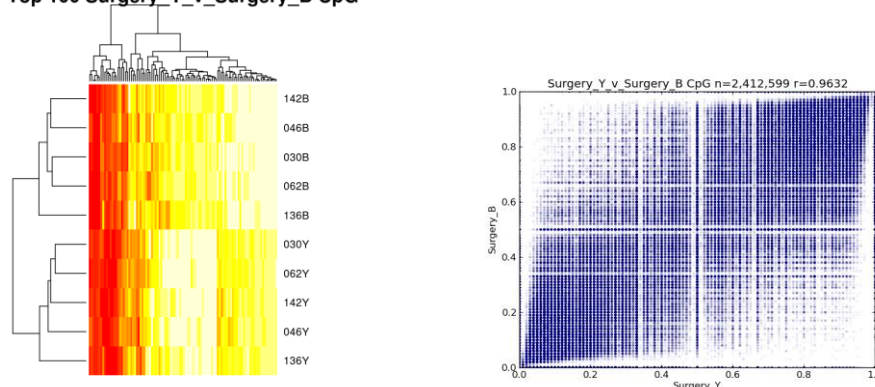
Group	Range Abs 260/280	Range Abs 260/230	Range DNA Concentration (ng/μl)
Control	1.88-2.28	1.21-2.59	28.5-70.2
Intervention	1.87-2.11	0.24-2.31	19.2-104.8
Surgery	1.91-2.04	0.37-2.19	23.4-96.2

One microgram of DNA from each sample was analyzed by Methyl-Mini Sequencing, which is a Reduced Repression Bisulfite Sequencing method that allows for high genome coverage and detection of 3-4 million CpG sites.

All samples had a bisulfite conversion rate of >98.25%; number of CpG (methylated) sites per sample ranged from 7,700,000-9,100,000, with minimum coverage of 5X per sample. A p-value <0.05 was significant. The top 2000 hypo-methylated (decreasing methylation) and hyper-methylated (increasing methylation) sites in the three groups were identified.

For LAGB patients, a heat map based on differences in methylation between baseline and one year (left) and a pairwise scatter plot of methylation patterns at baseline and one year (right) are shown below:

Top 100 Surgery_Y_v_Surgery_B CpG



Task #2: Profile metabolic activity in adipose tissue during surgical weight loss

During the year, 6 new patients undergoing laparoscopically placed adjustable gastric banding (LAGB) were enrolled in the study and provided baseline blood and adipose tissue samples. Blood samples were collected from LAGB patients at 83 follow-up examinations, ranging from the six-month follow-up to several years after LAGB. In all, 1238 aliquots of DNA, RNA, and plasma for biomarker assays were processed and stored. Staff collected 82 data sheets on LAGB patients and entered the data (BMI and other measures of adiposity) into our electronic database. Total RNA was isolated from 201 samples from 89 patients. The range of RNA concentrations was 3.8 ng/ μ l – 195.0 ng/ μ l. 184 samples were globin-cleared, fragmented, and run on U133A 2.0 expression arrays; call rates were 53.5-62.5%. An additional 12 samples were globin-cleared and fragmented.

Libraries were generated from 40 RNA samples from morbidly obese patients using the Illumina TruSeq RNA library prep v2 kit. QC checks using the Bio-analyzer and RT-PCR showed that the libraries were of good quality (average fragment length of 340 bp, no index dimers or SPRI bead carry over) and quantity (average concentration 48.7 nM, range 20 nM – 87 nM).

Using QC values, we calculated the amount of library needed to cluster a HiSeq flow cell using the Illumina cBot. We then normalized aliquots of the 40 RNA-seq libraries to 4 nM for pooling, denaturing, and further dilution prior to clustering. The 40 libraries were split into 8 groups of 5 libraries each, and within each group the 5 libraries were pooled and run together in a single lane. Libraries for lanes #1 – 3 were clustered at 18 pM, while libraries for lanes #4 – 8 were clustered at 15 pM based on previous QC data. After clustering the flow cell was ready for sequencing on the Illumina HiSeq platform.

The HiSeq was prepped for sequencing, which consisted of several wash steps, preparing and mixing reagents from various kits, fluidics checks, and sample sheet preparation. The clustered flow cell containing the 40 RNA-seq libraries was then run on a single-read flow cell, with single index identifier for 127 cycles (~5 days).

The run generated ~1.15 billion raw reads, equivalent to 138 billion bases sequenced. Cluster concentrations for lanes #4 – 8 were low at ~400k clusters per mm², lane #2 clustered best at 800k/mm². Quality of the clustering was very good with 94.5% of clusters passing filter and 98.7% of the reads identified. Quality of the reads was very good with 88% of the reads having a Q-score >30, which equates to 99.9% accuracy. All the data amounts to an average coverage for each sample of 49x, very close to our 50x objective.

Task #3: Use whole transcriptome analysis in the CRC to examine expression of previously identified genes

Nothing to report.

Task #4: Investigate gender and patient subgroup differences in molecular response

Nothing to report.

Task #5: Discover new genetic influences on heart disease by profiling micro-RNAs and rare RNA transcripts

Tasks #3, #4, and #5 are using current patients in the Cardiovascular Risk Clinic. No new participants entered the CRC program and no new blood samples were collected during the year. DNA samples (n=33) were assayed on the DMET™ SNP arrays and all samples had call rates >98%.

Task #6: Develop systems biology approach to integrate various types of risk factor data

Task #6 will utilize the large-scale DNA and RNA sequence data generated in Tasks #1 – #5, along with other CVD risk factor data collected in the Integrative Cardiac Health Program. To derive maximum information from the data, we are collaborating with scientists who have expertise in systems biology to integrate all of the different types of data. This approach will allow us to uncover inter-relationships and patterns within the data that may not be apparent when each modality is analyzed independently. Progress on this task will be made when we have sufficient DNA or RNA sequence data for analysis.

Discussion

The RNA concentrations resulting from isolation from blood or tissue continue to be sufficient for down-stream applications. The RIN numbers for all samples isolated for Tasks #2 – #5 indicated that all samples are high quality and should yield quality RNA sequence. Our first large-scale RNA sequencing run showed high quality clustering (94.5%) and <99.9% accuracy for 88% of the reads. During the next quarter, we plan to continue large-scale RNA sequencing.

Problems encountered during the year and plans to resolve them include:

1. Throughout most of the year, the HiSeq machine was not operational because air had entered the flow cells and could not be removed. This problem has been resolved, as Illumina service representatives recertified the HiSeq, MiSeq, and

cBot machines on June 11, 2014. The machines now have service contract coverage and all appear to be functioning normally.

2. The logistical issue of getting supplies continued throughout much of the year. We believe this issue has now been resolved and hope that supplies will be ordered and delivered in a timely manner.

4. KEY RESEARCH ACCOMPLISHMENTS:

1. Obtained first genome-wide methylation results and began interpretation
2. Generated first RNA sequence data on study participants

5. **CONCLUSION:** During the next year, our main focus will be on keeping all machines running and generating as much quality RNA sequencing and genome-wide DNA methylation data as possible. Once we have sufficient preliminary data, we will begin integrating the various types of data and begin preparing abstracts for presentation and publication.

6. **PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS:** Nothing to report.

7. **INVENTIONS, PATENTS AND LICENSES:** Nothing to report.

8. **REPORTABLE OUTCOMES:** Nothing to report.

9. **OTHER ACHIEVEMENTS:** Nothing to report.

10. **REFERENCES:** Nothing to report.

11. **APPENDICES:** Nothing to report.